

# Visualizing AMIA : A Medical Informatics Knowledge Domain Analysis

Marie Synnestvedt MEd,<sup>1,2</sup> Chaomei Chen PhD<sup>2</sup>

<sup>1</sup>Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, Philadelphia PA

<sup>2</sup>College of Information Science and Technology, Drexel University, Philadelphia PA

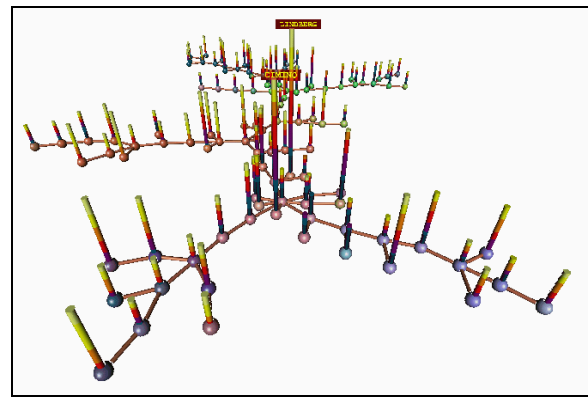
**Abstract.** Medical Informatics has been described as having a “long and delayed adolescence” which continues to “find itself in search of self-definition”,<sup>1</sup> and the AMIA Symposium Proceedings have been viewed as an indicator of trends in the field<sup>2</sup>. This pilot study investigated the feasibility of applying a knowledge domain visualization approach<sup>3</sup> to clarifying the domain of medical informatics based on the AMIA publications. Document co-citation analysis (DCA) is combined with Pathfinder Network Scaling (PFNET), visualization, and animation to develop a 3-D knowledge landscape.

**Background.** Previous studies using a literature based approach to reveal the structure of Medical Informatics have examined journal cocitation patterns<sup>4</sup> and co-occurrence of indexing terms<sup>5</sup> among a core journal set identified from Institute for Scientific Information (ISI) Web of Science online databases covering the Indexing period January 1993-July 1995. JAMIA did not begin publication until 1994, and ISI has recently entered citations from all of the Proceedings papers published since 1994 into Web of Science<sup>2</sup>. The unique opportunity that ISI indexing offers is the availability of data on cited references for DCA.

**Methods.** Bibliographic data on 1,986 JAMIA and AMIA Symposium papers and their cited references for the 1994-2002 indexing period were obtained from Web of Science. Articles that were cited more than five times were selected for DCA, and cocitation counts for those selected articles transformed into a matrix for factor analysis. The factors identified through Principle Component Analysis (PCA) are considered to correspond to groupings of research topics and specialties within scientific disciplines<sup>3</sup>. The DCA network was submitted to Pathfinder network scaling, and the resultant PFNET rendered as a virtual landscape of the knowledge domain. In PFNETs, nodes represent articles, the links represent salient cocitation paths. Dominant articles can be seen as those with relatively many links to other articles. Links between articles and dominant articles define specialties, and links between dominant articles connect specialties<sup>6</sup>. A citation time series has also been color mapped to a stacked bar on top of

each document's node (Figure 1).

**Results.** The top loading factor from PCA accounted for 42% of the variance, and a review of the top 20 papers associated with that factor shows an area of specialization related to controlled terminologies and vocabularies.



**Figure 1. Knowledge Landscape**

**Conclusions.** There is now sufficient cocitation data available to apply this type of knowledge domain analysis to the AMIA publications. Future work will enhance the mapping by adding citation data from 1996-2002 Symposium abstracts.

## References

1. Kohane IS. Preface. Proceedings/AMIA Annual Symposium. 2002:viii.
2. Lindberg DA. Humphreys BL. "You have to be there": twenty-five years of SCAMC/AMIA symposia. JAMIA. 2002;9(4):332-45.
3. Chen C. Paul RJ. O'Keefe B. Fitting the jigsaw of citations: Information visualization in domain analysis. JASIST. 2001;52(4), 315-330.
4. Morris TA. McCain KW. The structure of medical informatics journal literature. JAMIA. 1998;5(5):448-66.
5. Morris TA. Structural relationships within medical informatics. Proceedings/AMIA Annual Symposium. 2000:590-4.
6. White HD. Pathfinder Networks and Author Cocitation Analysis: A Remapping of Paradigmatic Information Scientists. JASIST. 2003;54(5):423-434.